Visione Artificiale

(a.a. 2003/04)

Prof. Marcello Pelillo

pelillo@dsi.unive.it

http://www.dsi.unive.it/~pelillo

Cosa significa "vedere"?

Ottenere una descrizione del mondo (3-D) da una o piu' immagini (2-D)

Visione artificiale (o computazionale):

Realizzare una macchina che "veda"

(o un algoritmo che consenta ad una "macchina di "vedere")

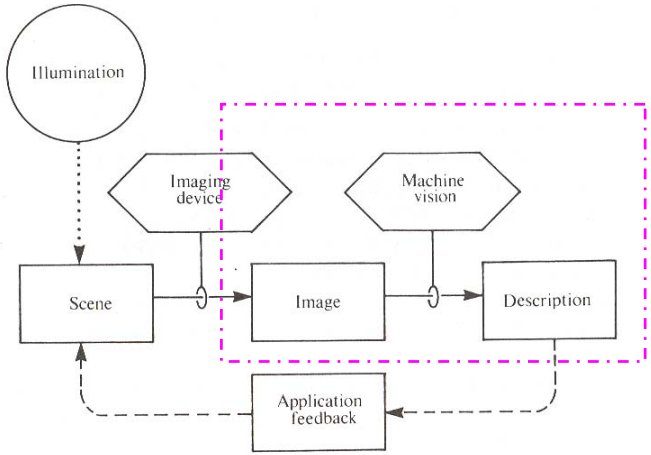
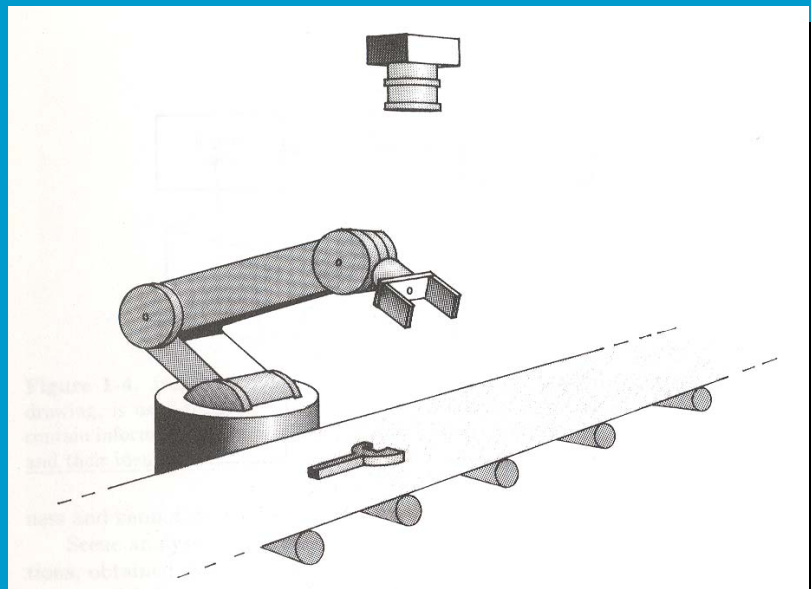


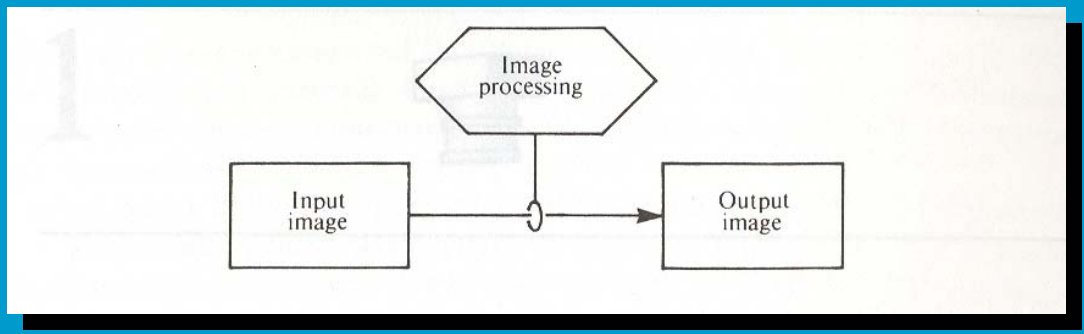
Figure 1-2. The purpose of a machine vision system is to produce a symbolic description of what is being imaged. This description may then be used to direct the interaction of a robotic system with its environment. In some sense, the vision system's task can be viewed as an inversion of the imaging process.

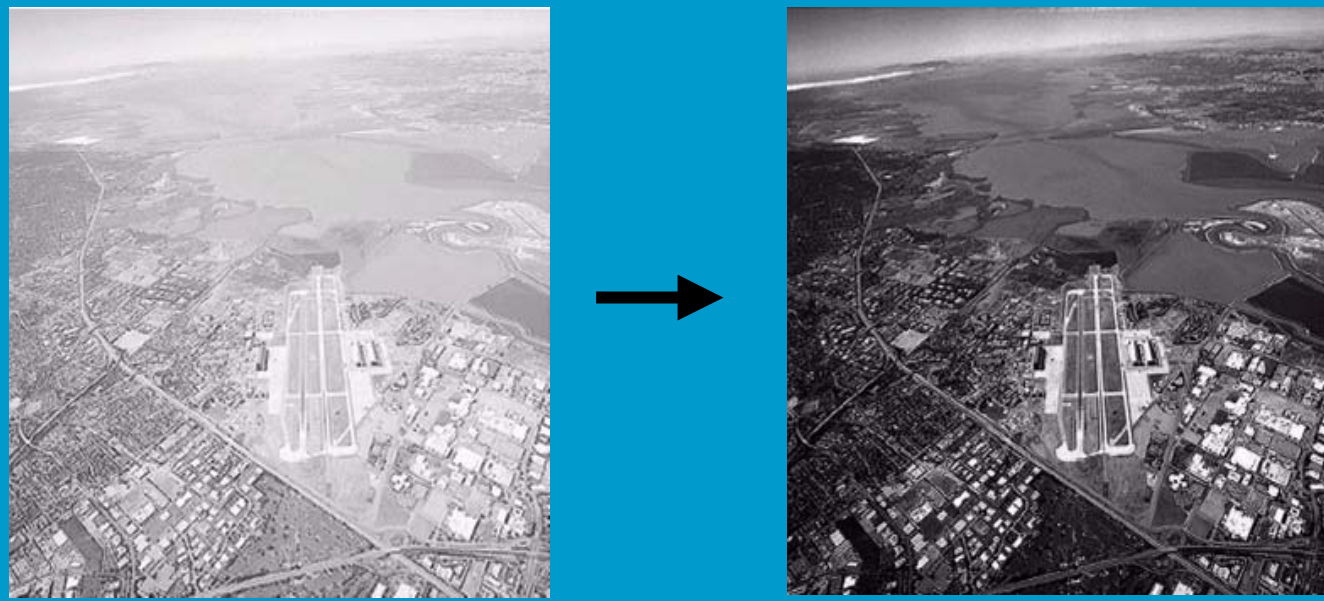


Discipline affini

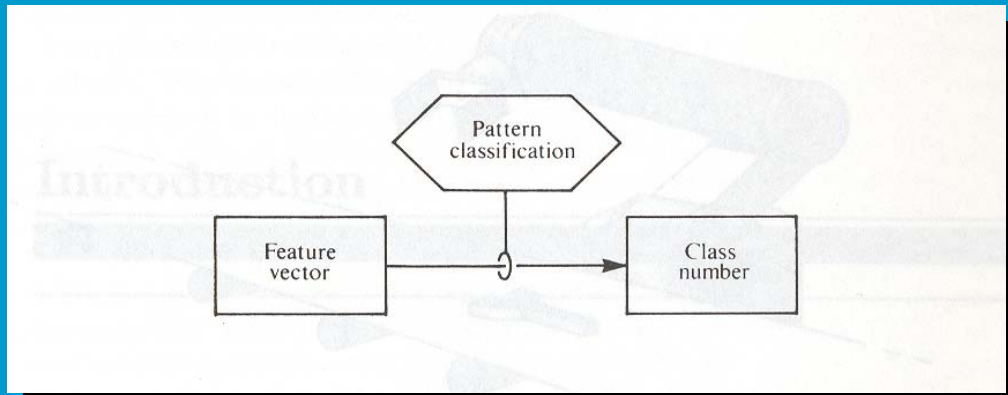
- Elaborazione delle immagini
- Pattern recognition
- Analisi delle scene

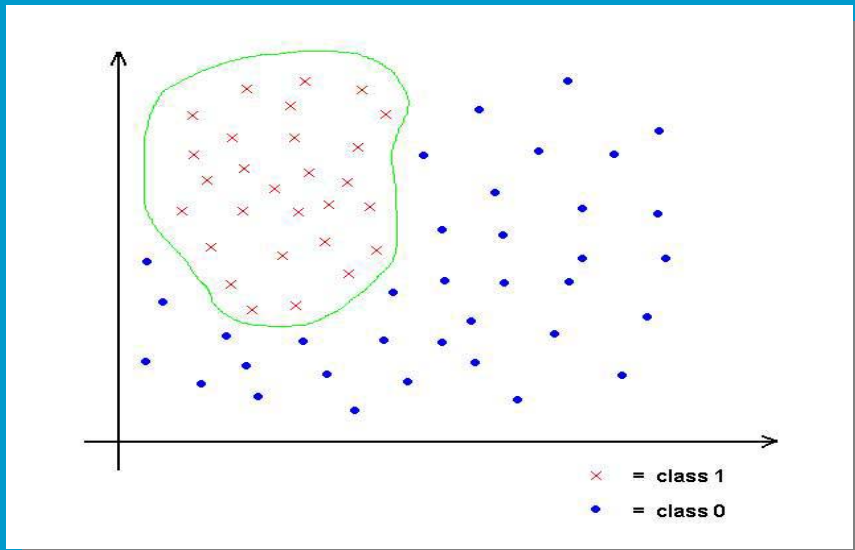
Elaborazione delle immagini



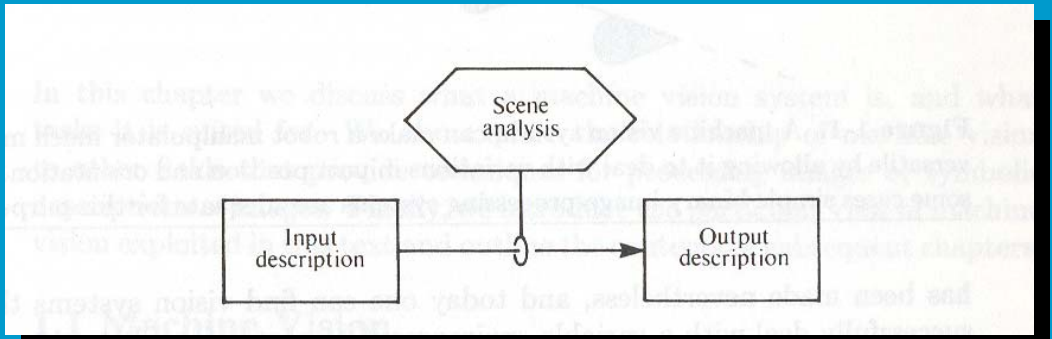


Pattern recognition





Scene analysis



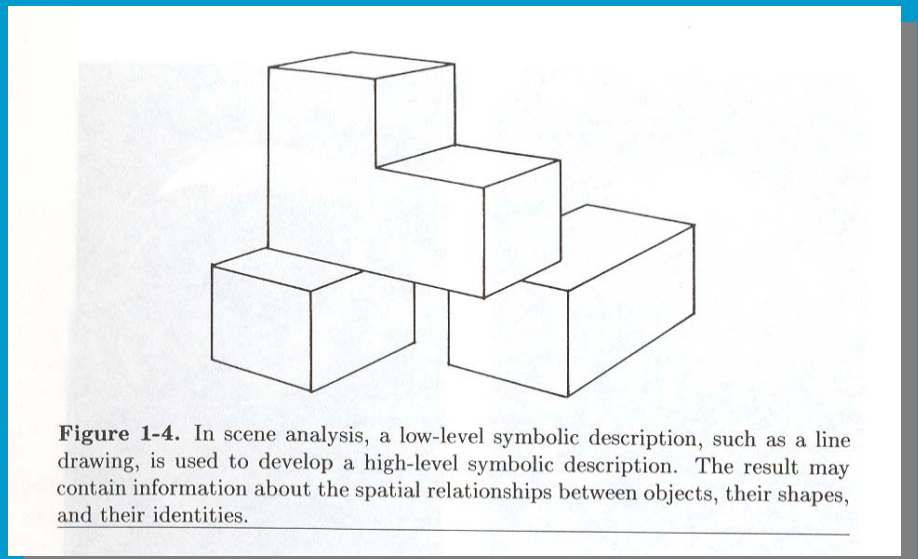
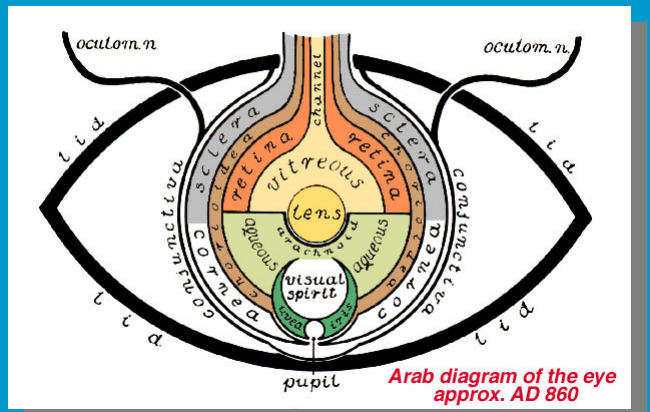


Figure 1-4. In scene analysis, a low-level symbolic description, such as a line drawing, is used to develop a high-level symbolic description. The result may contain information about the spatial relationships between objects, their shapes, and their identities.

Teorie della visione: Cenni storici



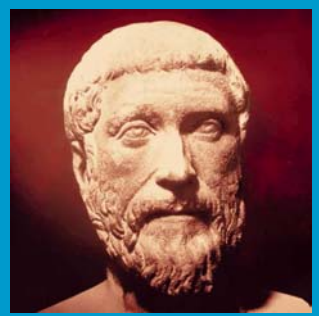
Arab diagram of the eye
approx. AD 860

La teoria emissionista (visione "tattile")

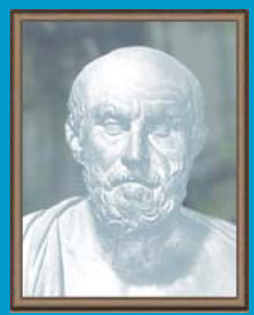
L'occhio emette un fascio di raggi che, viaggiando nello spazio vanno a urtare gli oggetti. L'urto tra il raggio visivo e l'oggetto suscita la sensazione della visione.

Principali esponenti

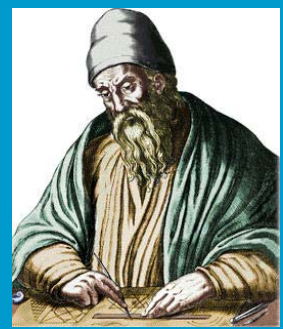
Pitagora



Empedocle
(490-430 a. C.)



Euclide

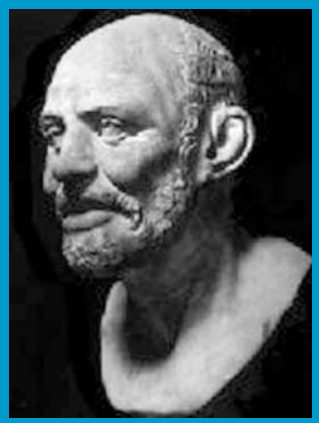


La teoria intromissionista

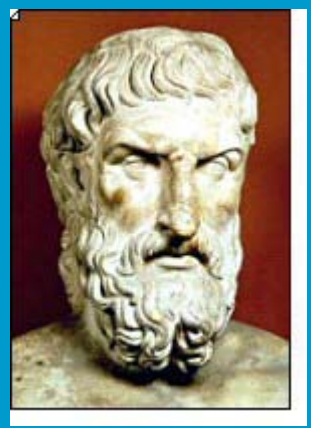
Gli oggetti inviano continuamente nello spazio ad essi circostante le immagini di se stessi. Queste immagini (dette eidola) entrano nell'occhio attraverso la pupilla, cosi' rivelandosi.

Principali esponenti

Democrito
(460-360 a.C.)



Epicuro



Lucrezio

Una combinazione delle due

Platone (428-347 a.C.), in vari dialoghi (Timeo, Menone, Teeteto) propone una teoria altrenativa che combina le due precedenti.

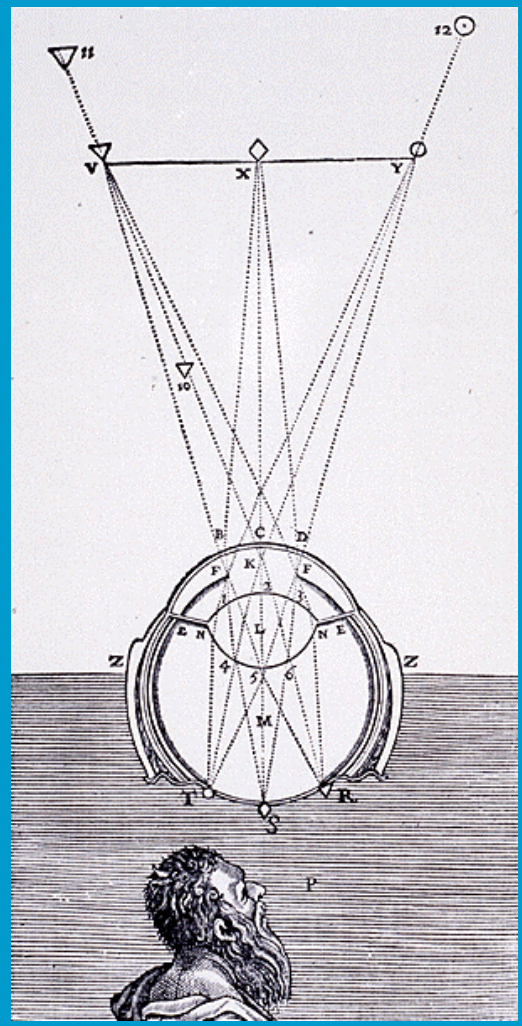
ibn al-Haytham (Alhazen) (965-1039 d.C.)

Alhazen demoli' definitavamente la teoria emissionista:

•Persistenza delle immagini retiniche ("after-image")

Keplero e la teoria dell'immagine retinica

Nel 1604 Keplero pose fine alla disputa sviluppando la moderna teoria delle immagini retiniche.



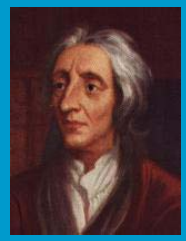
Nativismo contro Empirismo

Cartesio

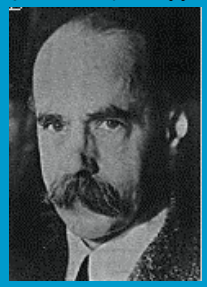
Kant

Berkeley

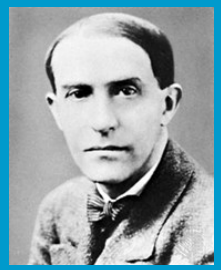
Locke



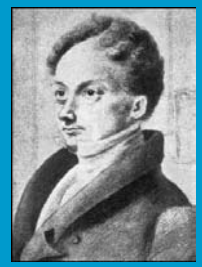
Wertheimer



Koehler

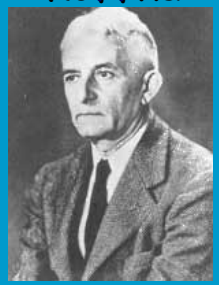


Mill



Hume

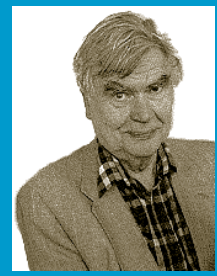
Koffka



Kanizsa

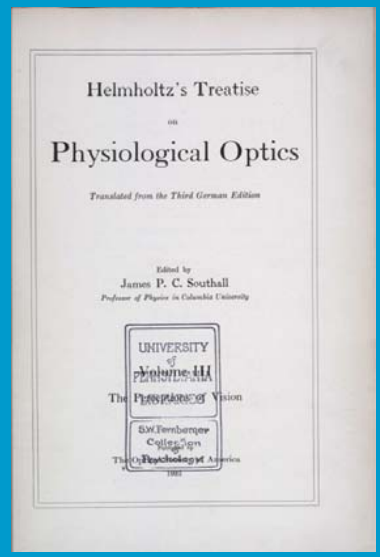
Helmholtz

Gregory



La percezione come "inferenza inconscia": Helmholtz

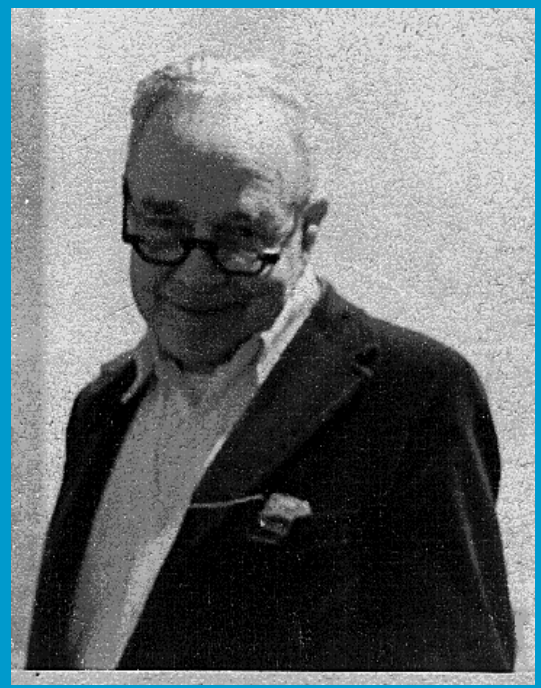
The psychic activities that lead us to infer that there in front of us at a certain place there is a certain object of a certain character, are generally not conscious activities, but unconscious ones. In their result they are equivalent to a *conclusion*, to the extent that the observed action on our senses enables us to form an idea as to the possible cause of this action; although, as a matter of fact, it is invariably simply the nervous stimulations that are perceived directly, that is, the actions, but never the external objects themselves. But what seems to differentiate them from a conclusion, in the ordinary sense of that word, is that a conclusion is an act of conscious thought. An astronomer, for example, comes to real conscious conclusions of this sort, when he computes the positions of the stars in space, their distances, etc., from the perspective images he has had of them at various times and as they are seen from different parts of the orbit of the earth. His conclusions are based on a conscious knowledge of the laws of optics. In the ordinary acts of vision this knowledge of optics is lacking. Still it may be permissible to speak of the psychic acts of ordinary perception as *unconscious conclusions*, thereby making a distinction of some sort between them and the common so-called conscious conclusions. And while it is true that there has been, and probably always will be, a measure of doubt as to the similarity of the psychic activity in the two cases, there can be no doubt as to the similarity between the results of such unconscious conclusions and those of conscious conclusions.



James J. Gibson (1904-1979) e l'approccio ecologico

"The belief of the empiricists that the perceived meanings and values of things are supplied from past the experience of the observer will not do. But even worse is the belief of nativists that meanings and values are supplied from the past experience of the race by innate ideas."

(Gibson, 1979)



David Marr e l'approccio computazionale

I tre livelli di elaborazione:

- Computazionale

- Algoritmico

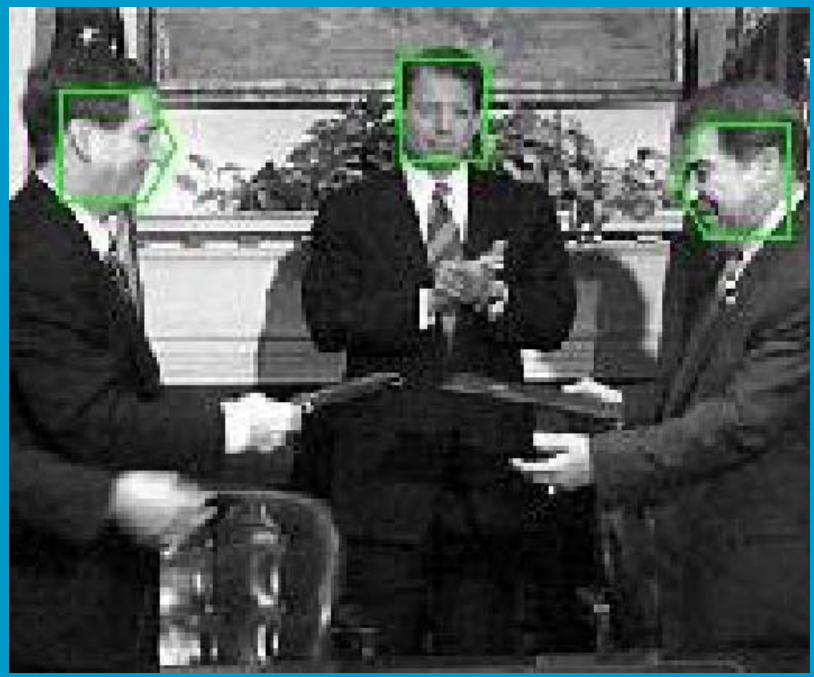
- Implementativo (hardware)

Alcune applicazioni della visione artificiale

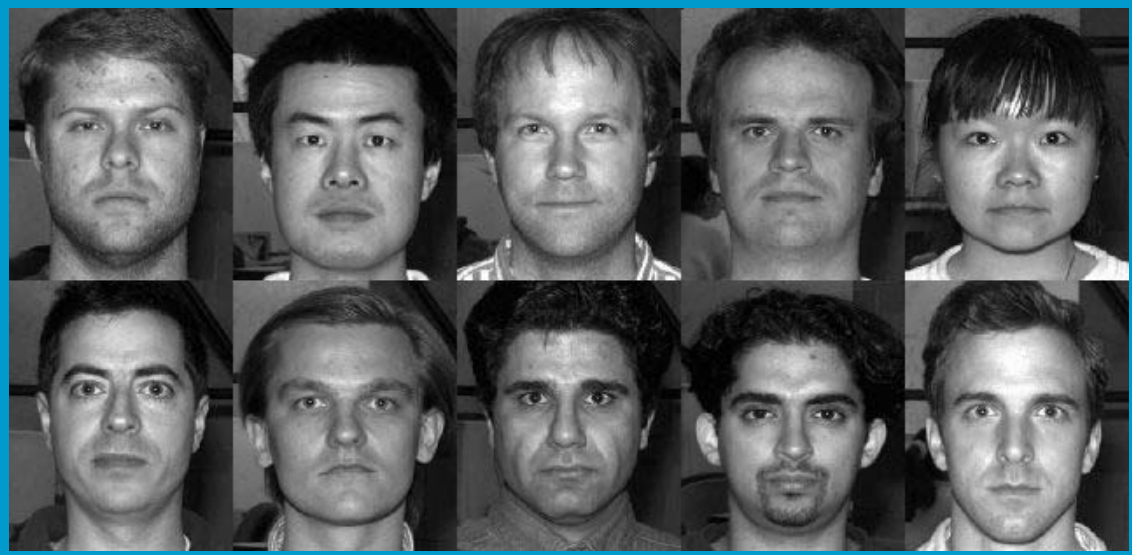
Ricerca in database di immagini

From a search for horse pix in 100 horse images and 1086 non-horse images

Rilevamento di volti (face detection)



Riconoscimento di volti (face recognition)



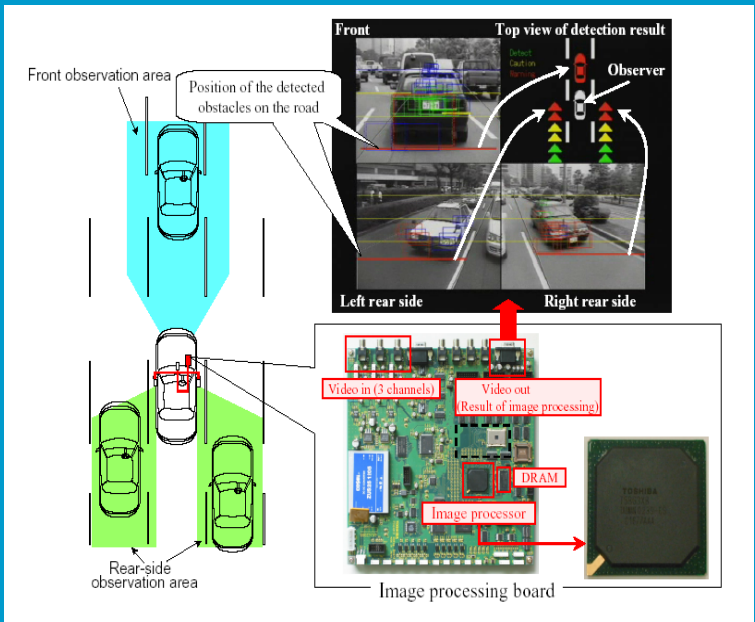
Acquisizione di volti (Yale)

Robotica

Il robot SONY SDR-4X II

Sistemi di sorveglianza per automobili

TOSHIBA



Veicoli autonomi

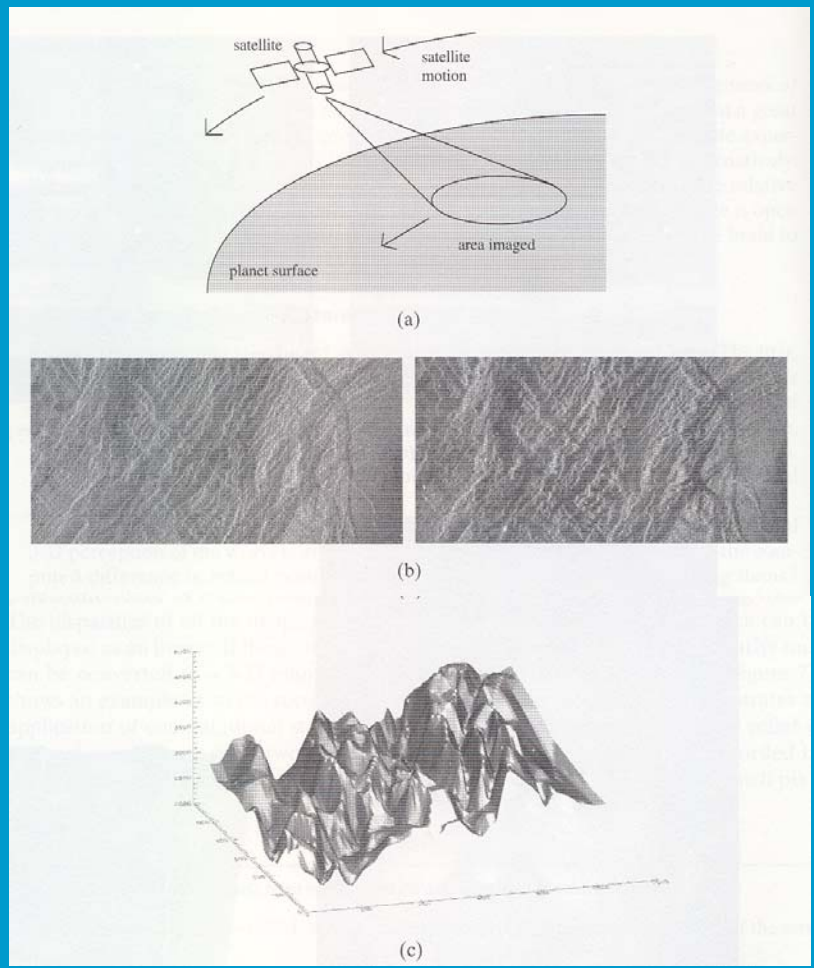
ARGO
(Universita' di Parma)

TG1, 16 Settembre '99

Rilevamenti da satellite

Ricostruzione stereo della superfice di Venere da una coppia di immagini satellitari.

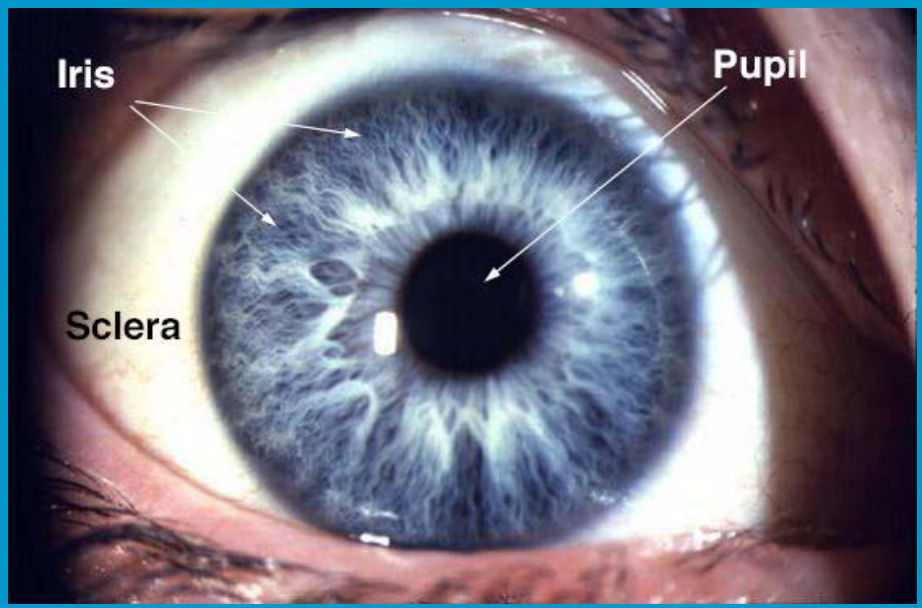
(Institute for Computer Graphics and Vision, Technical University of Graz, Austria)

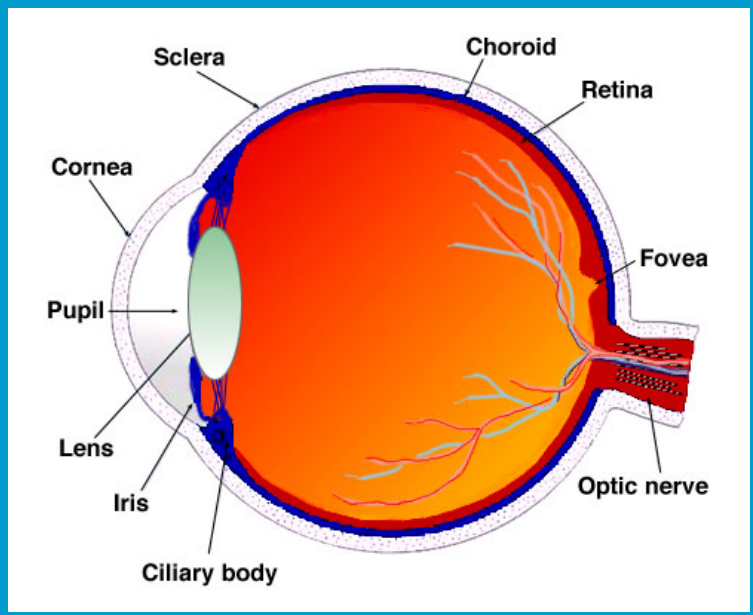


Fisiologia della visione (cenni)

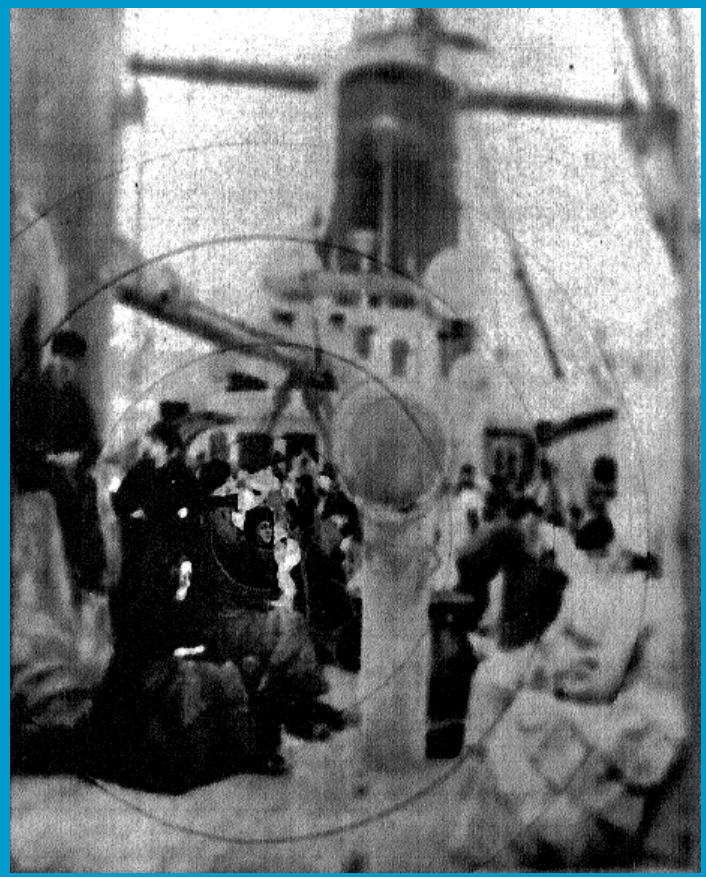
Imitare la natura?

L'occhio


Iris
Pupil
Sclera


Sclera
Choroid
Retina
Cornea
Fovea
Pupil
Lens
Iris
Optic nerve
Ciliary body

Visione foveale



La retina

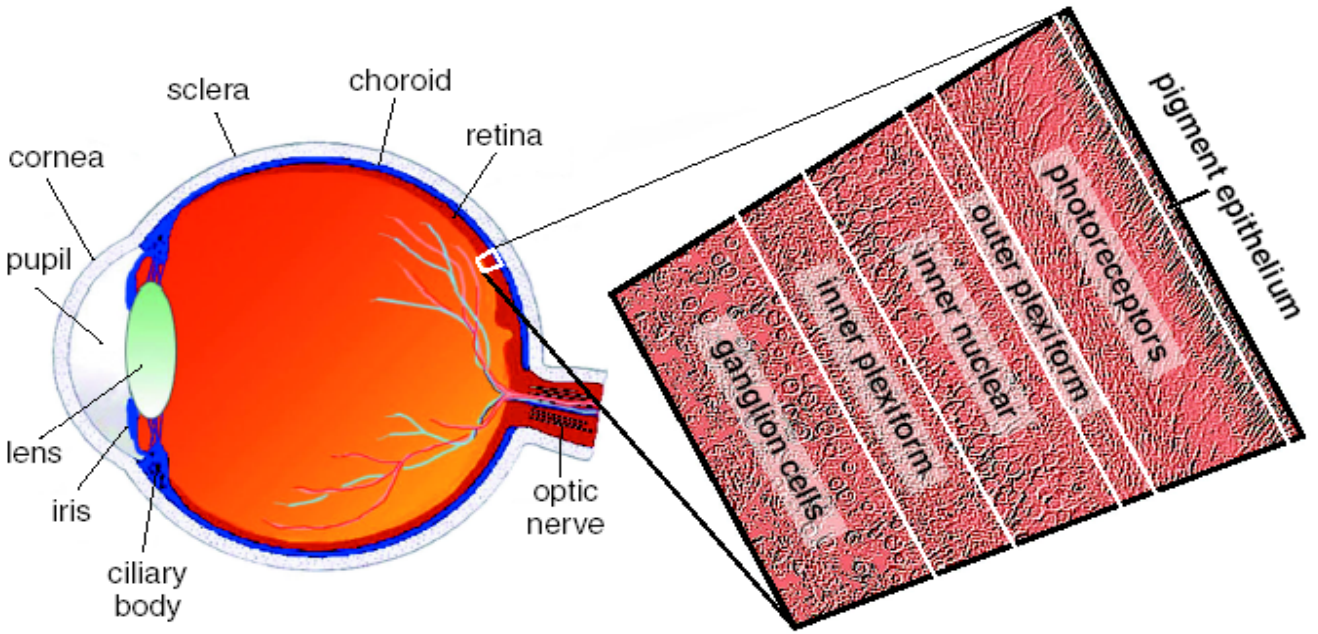


Figure 2. Diagram of a human eye shows its various structures *(left)*. A thin piece of retina is enlarged in a photomicrograph *(right)*, revealing its layers. The photoreceptors lie against a dark row of cells called the pigment epithelium. (Drawing by the author. Except where noted, photographs by Nicolas Cuenca and the author.)

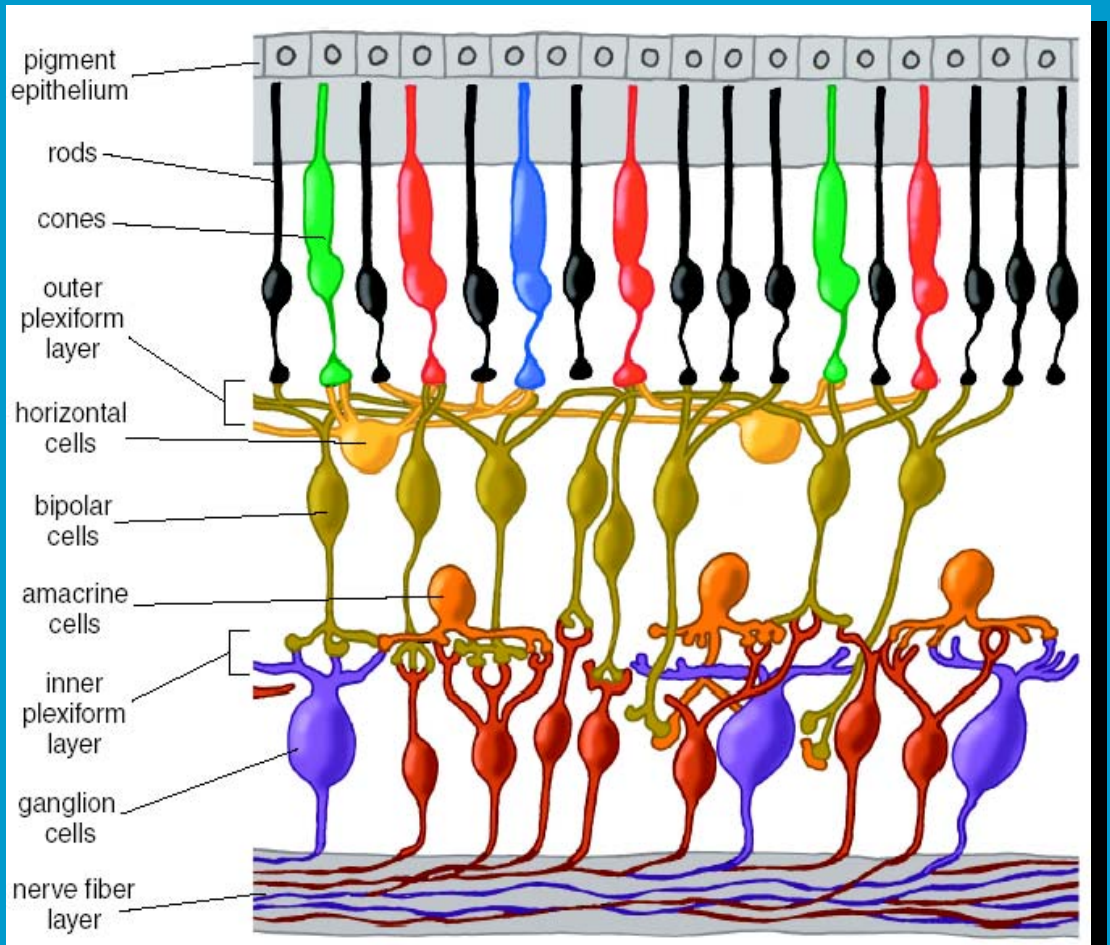
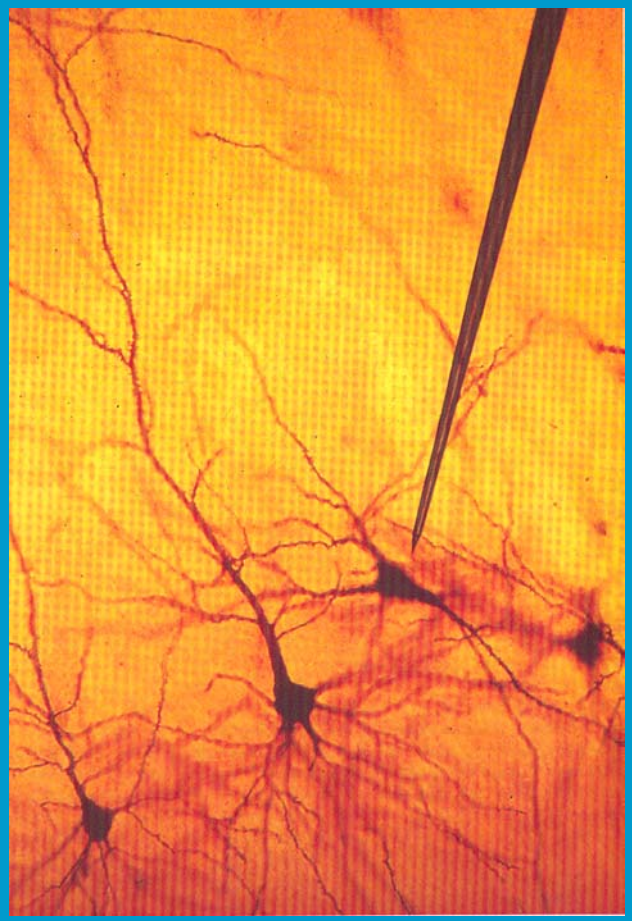
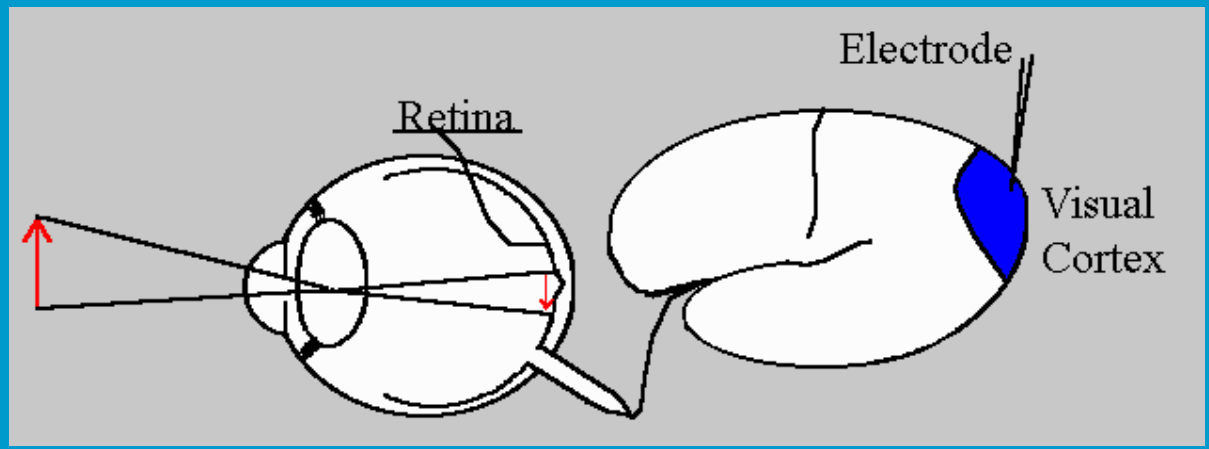


Figure 3. Cells in the retina are arrayed in discrete layers. The photoreceptors are at the top of this rendering, close to the pigment epithelium. The bodies of horizontal cells and bipolar cells compose the inner nuclear layer. Amacrine cells lie close to ganglion cells near the surface of the retina. Axon-to-dendrite neural connections make up the plexiform layers separating rows of cell bodies.

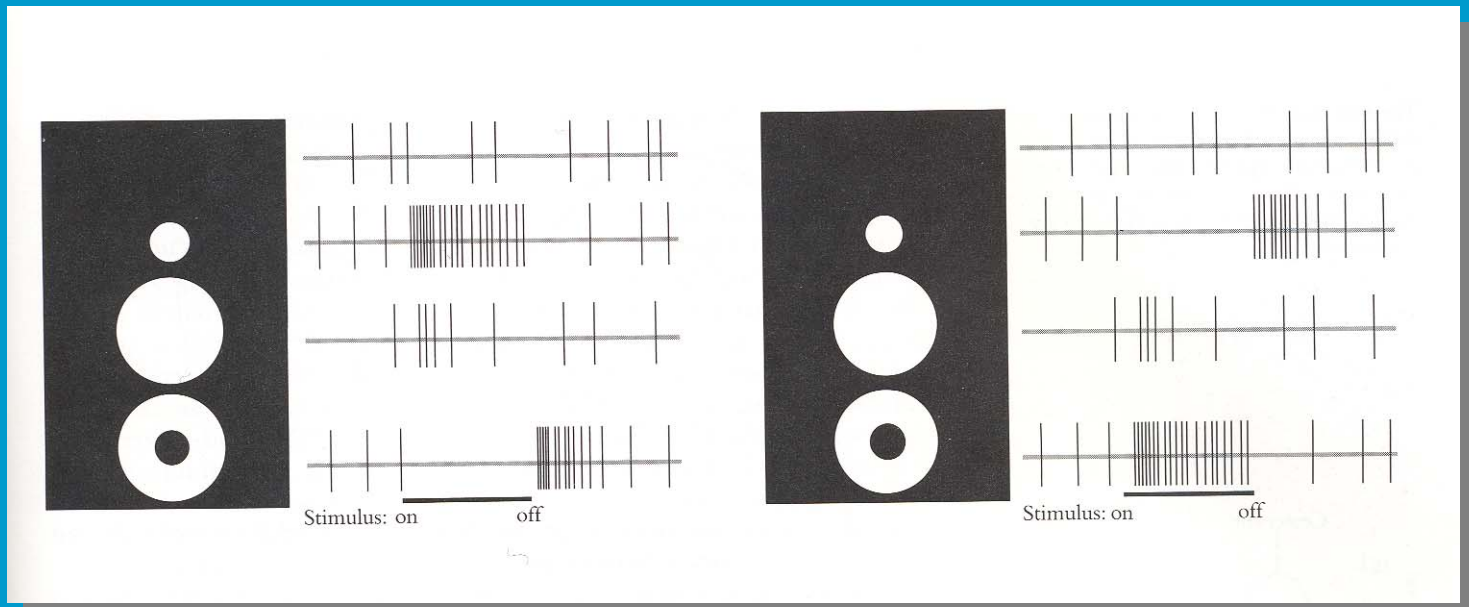
Registrazioni cellulari



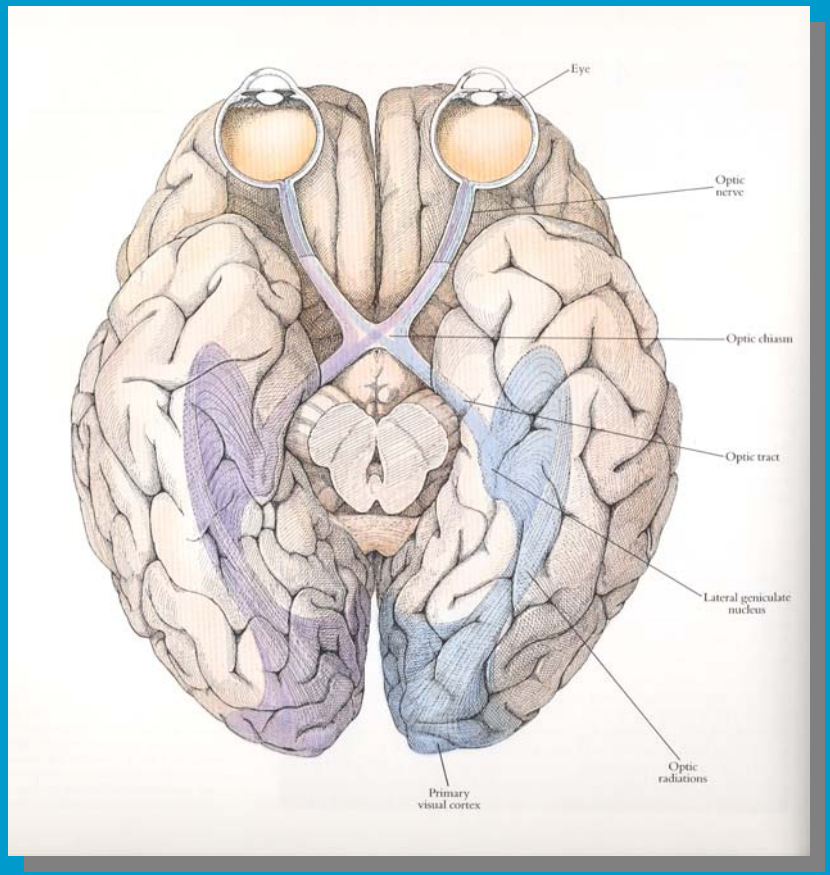
The visual cortex in a monkey, stained by the Golgi method, shows a few pyramidal cells—a tiny fraction of the total number in such a section. The entire height of the photograph represents about 1 millimeter. A tungsten microelectrode, typical of what is used for extracellular recordings, has been superimposed, to the same scale.

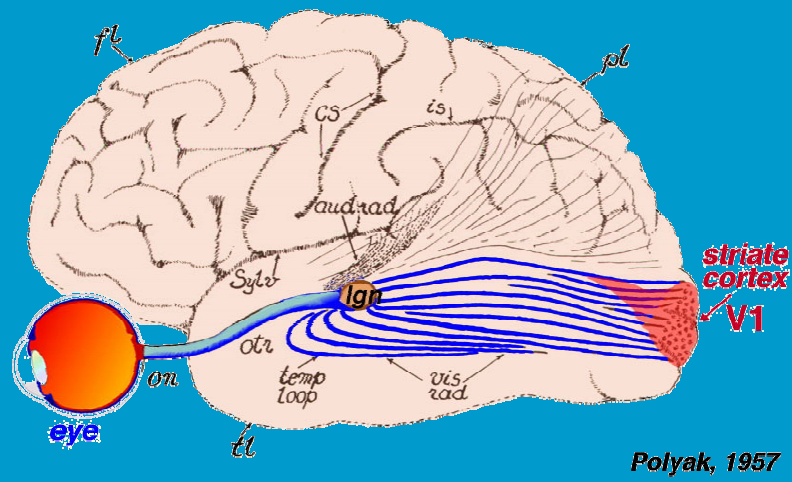


L'ouput della retina

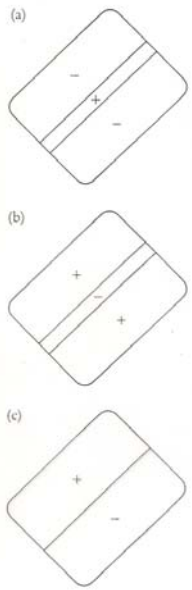


Oltre la retina

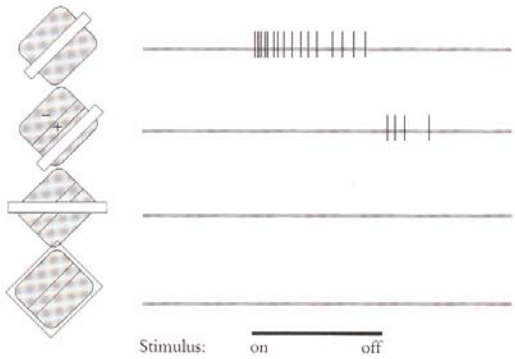




Cellule semplici

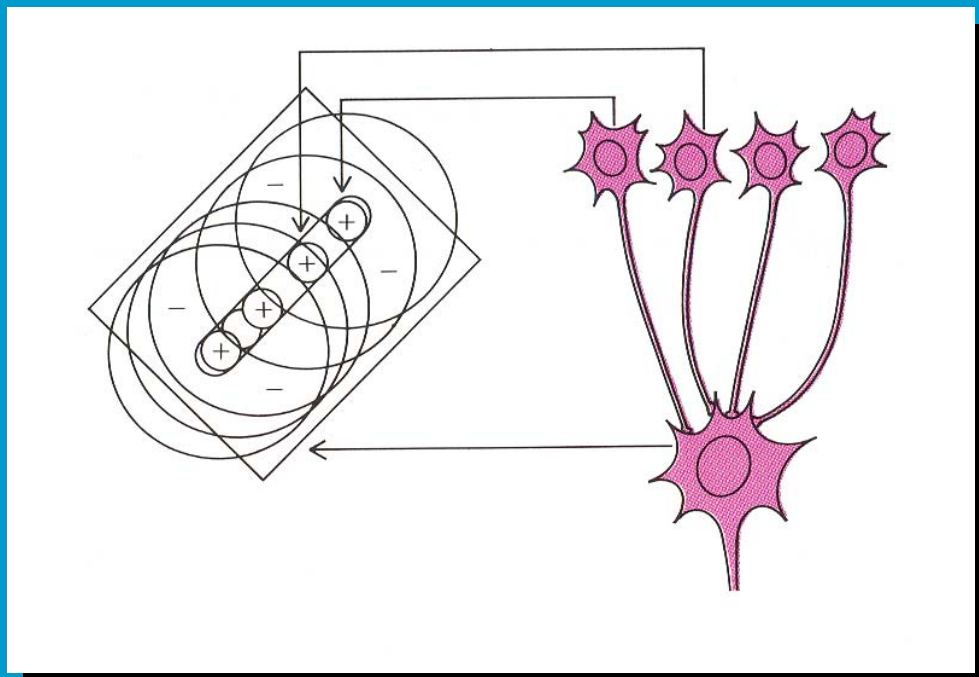


Three typical receptive-field maps for simple cells. The effective stimuli for these cells are (a) a slit covering the plus (+) region, (b) a dark line covering the minus (−) region, and (c) a light-dark edge falling on the boundary between plus and minus.



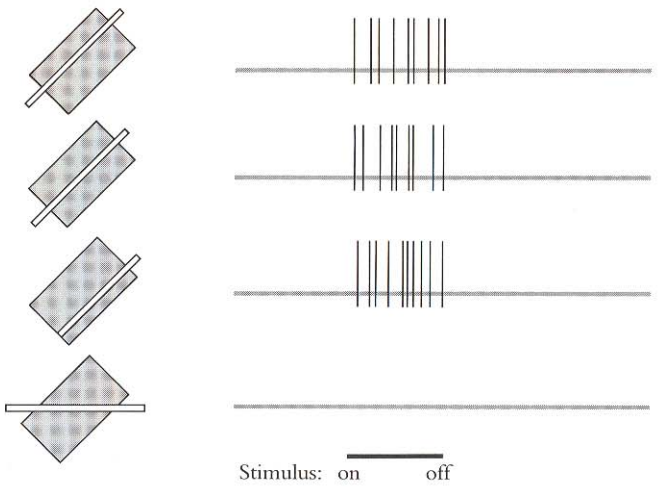
Various stimulus geometries evoke different responses in a cell with receptive field of the type in diagram a of the previous figure. The stimulus line at the bottom indicates when the slit is turned on and, 1 second later, turned off. The top record shows the response to a slit of optimum size, position, and orientation. In the second record, the same slit covers only part of an inhibitory area. (Because this cell has no spontaneous activity to suppress, only an off discharge is seen.) In the third record, the slit is oriented so as to cover only a small part of the excitatory region and a proportionally small part of the inhibitory region; the cell fails to respond. In the bottom record, the whole receptive field is illuminated; again, there is no response.

Dalle gangliari alle semplici



This type of wiring could produce a simple-cell receptive field. On the right, four cells are shown making excitatory synaptic connections with a cell of higher order. Each of the lower-order cells has a radially symmetric receptive field with on-center and off-surround, illustrated by the left side of the diagram. The centers of these fields lie along a line. If we suppose that many more than four center-surround cells are connected with the simple cell, all with their field centers overlapped along this line, the receptive field of the simple cell will consist of a long, narrow excitatory region with inhibitory flanks. Avoiding receptive-field terminology, we can say that stimulating with a small spot anywhere in this long, narrow rectangle will strongly activate one or a few of the center-surround cells and in turn excite the simple cell, although only weakly. Stimulating with a long, narrow slit will activate all the center-surround cells, producing a strong response in the simple cell.

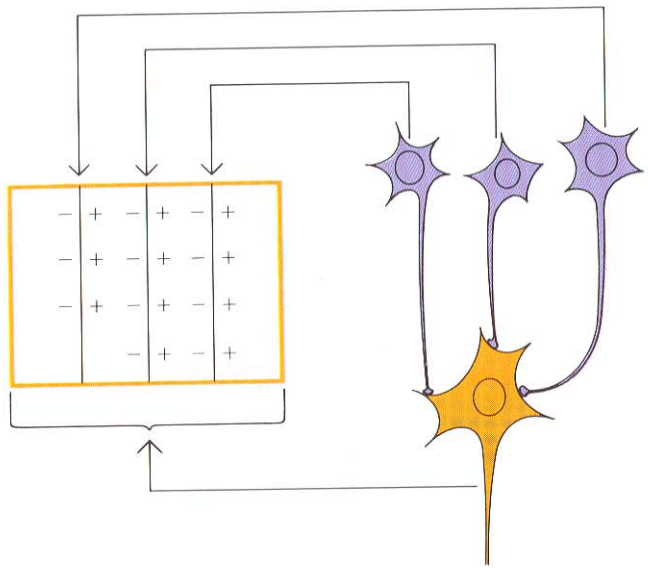
Cellule complesse



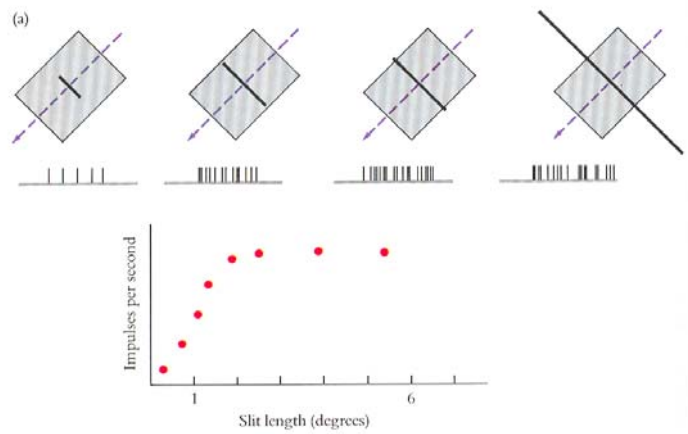
A long, narrow slit of light evokes a response wherever it is placed within the receptive field (rectangle) of a complex cell, provided the orientation is correct (upper three records). A nonoptimal orientation gives a weaker response or none at all (lower record).

Dalle semplici alle complesse

This wiring diagram would account for the properties of a complex cell. As in the figure on page 74, we suppose that a large number of simple cells (only three are shown here) make excitatory synapses with a single complex cell. Each simple cell responds optimally to a vertically oriented edge with light to the right, and the receptive fields are scattered in overlapping fashion throughout the rectangle. An edge falling anywhere within the rectangle evokes a response from a few simple cells, and this in turn evokes a response in the complex cell. Because there is adaptation at the synapses, only a moving stimulus will keep up a steady bombardment of the complex cell.



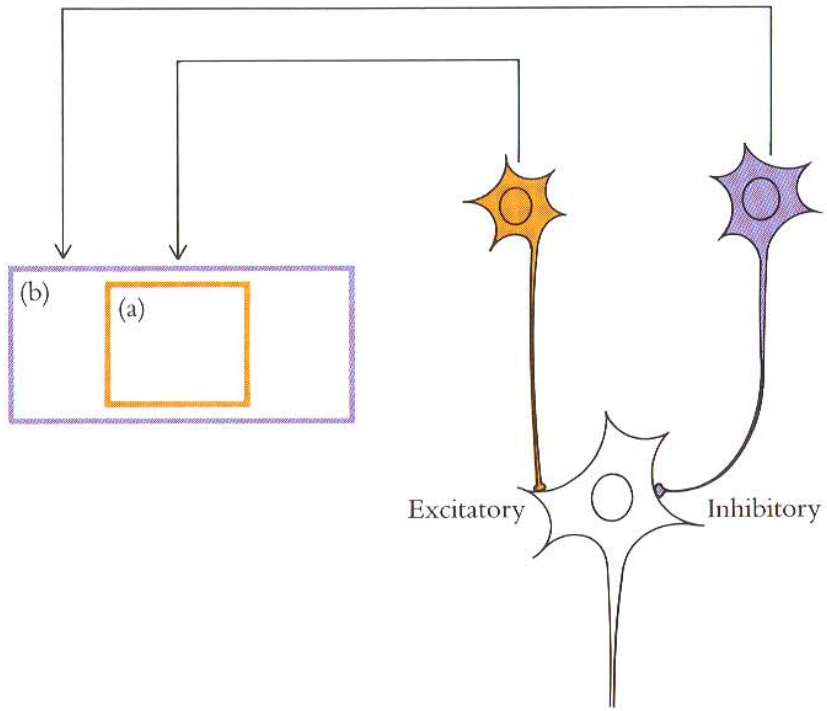
Cellule "end-stopped"



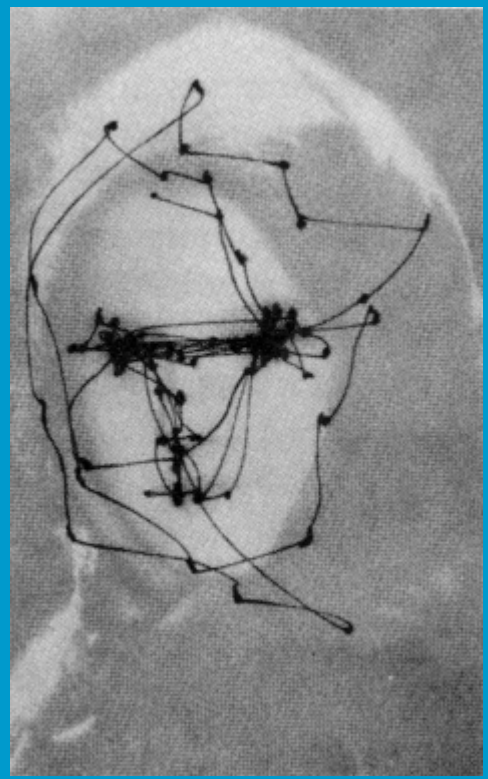
Top: An ordinary complex cell responds to various lengths of a slit of light. The duration of each record is 2 seconds. As indicated by the graph of response versus slit length, for this cell the response increases with length up to about 2 degrees, after which there is no change. *Bottom:* For this end-stopped cell, responses improve up to 2 degrees but then decline, so that a line 6 degrees or longer gives no response.

Un modello per l'end-stopping

In an alternative scheme, one cell does the inhibiting, a cell whose receptive field covers the entire area, (a) plus (b) and (c). For this to work, we have to assume that the inhibiting cell responds only weakly to a short slit when (a) is stimulated, but responds strongly to a long slit.



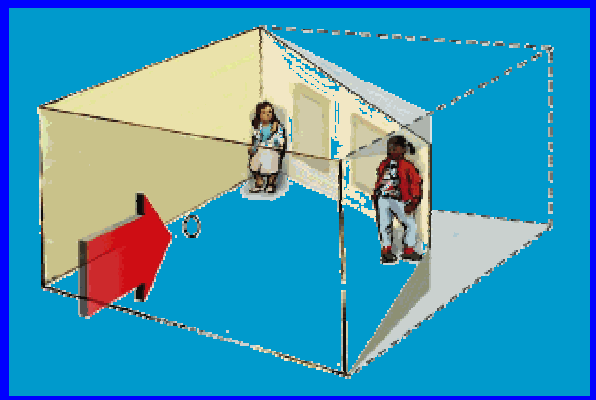
Il movimento degli occhi



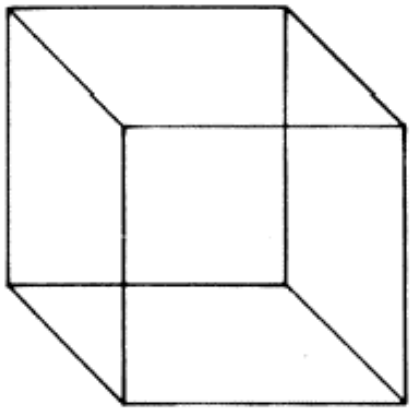
Ambiguita', inconsistenze, illusioni

La stanza di Ames

The Ames Room



Il cubo di Necker



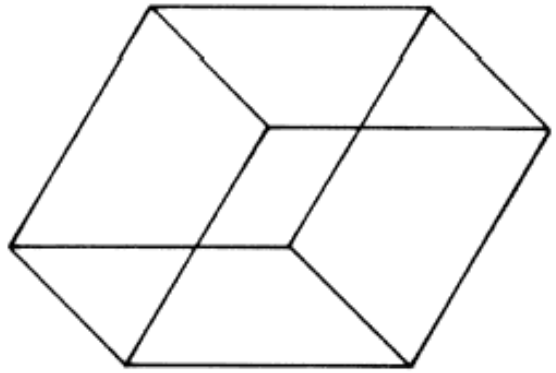
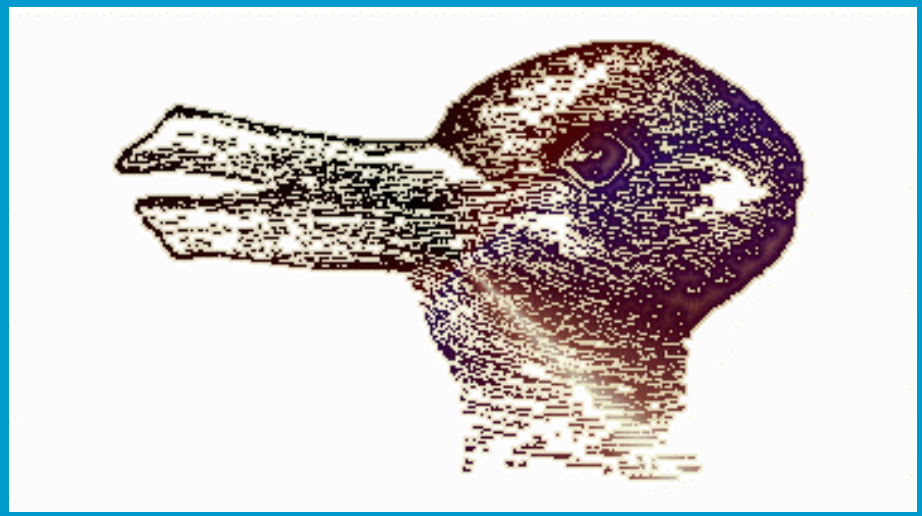


FIG. 2 (a) Necker cube. This is the most famous of many depth-ambiguous, figures. (When presented with no background it changes in shape with each reversal, the apparent back being larger than the apparent front face.) (b) Necker rhomboid. This is the original form, presented by L. A. Necker in 1832.

Altre immagini bi-stabili



Illusione di Mueller-Lyer

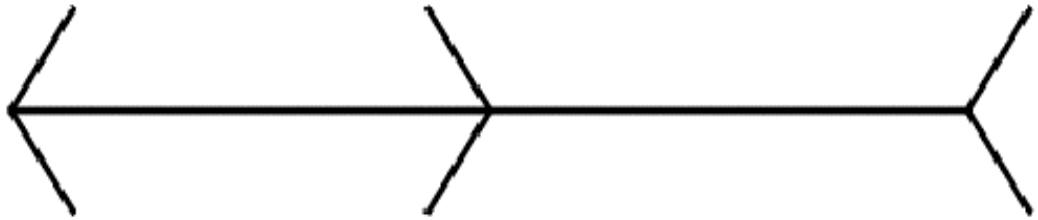


FIG. 5 Müller-Lyer arrows figure 1889. The most famous illusion: the outward-going 'arrow heads' produce expansion of the 'shaft' and the inward-going heads contraction.

Illusione di Ponzo

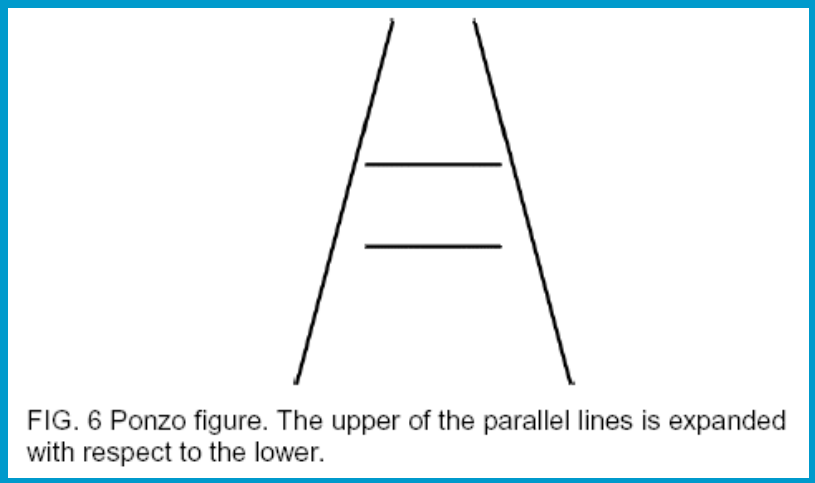
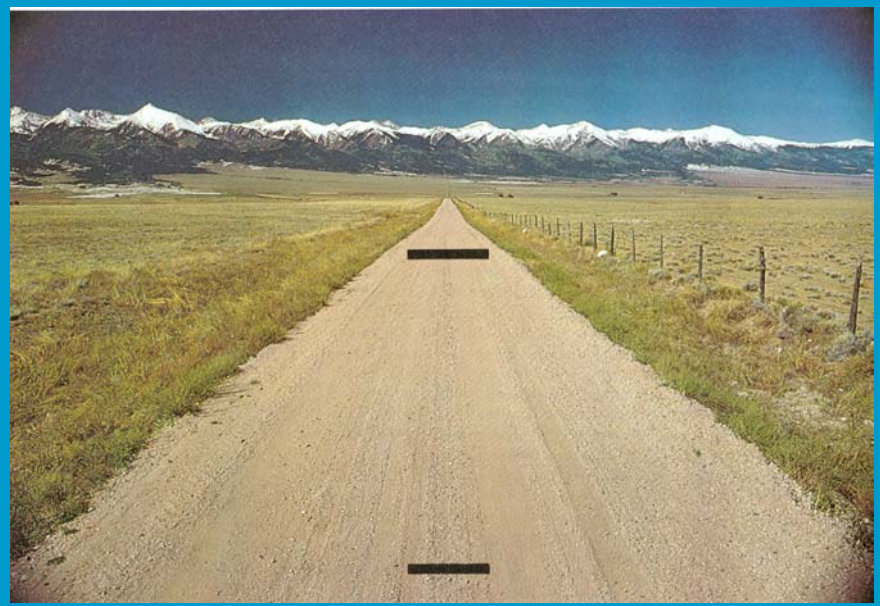
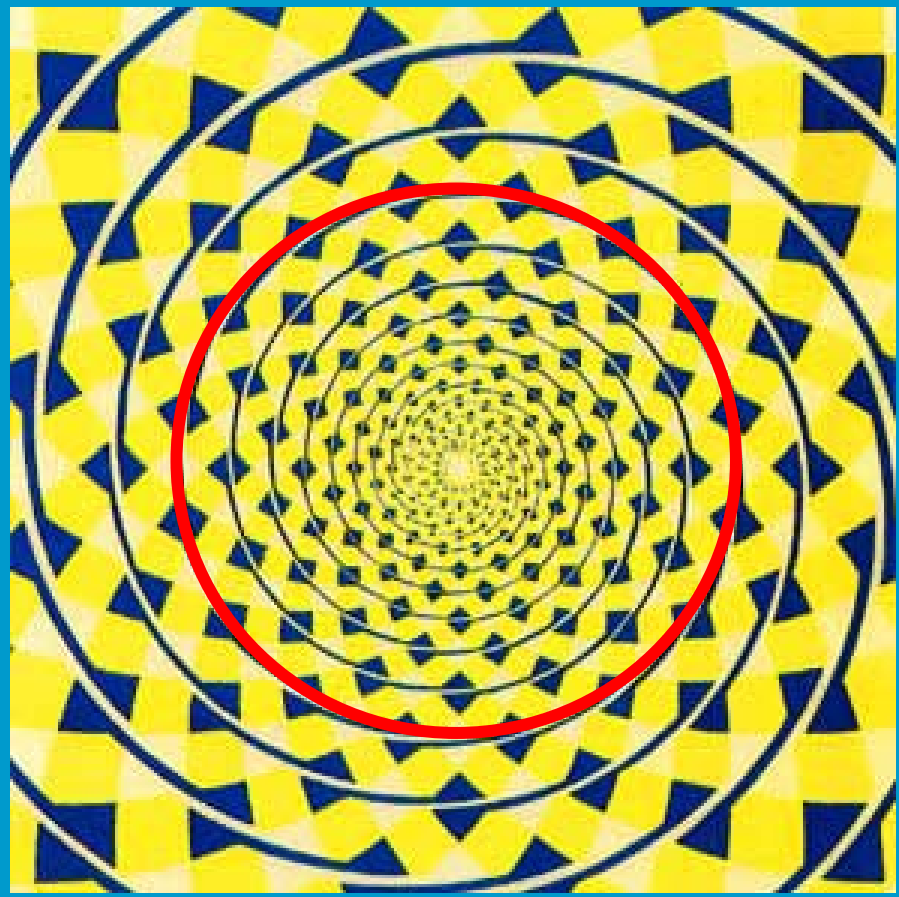


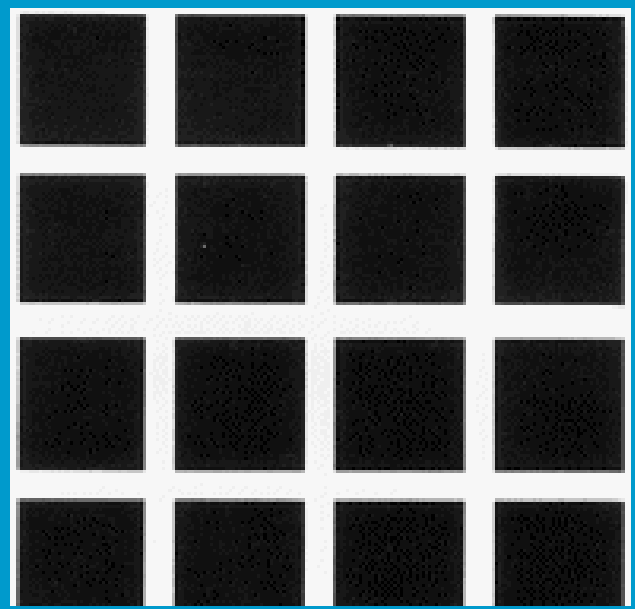
FIG. 6 Ponzo figure. The upper of the parallel lines is expanded with respect to the lower.



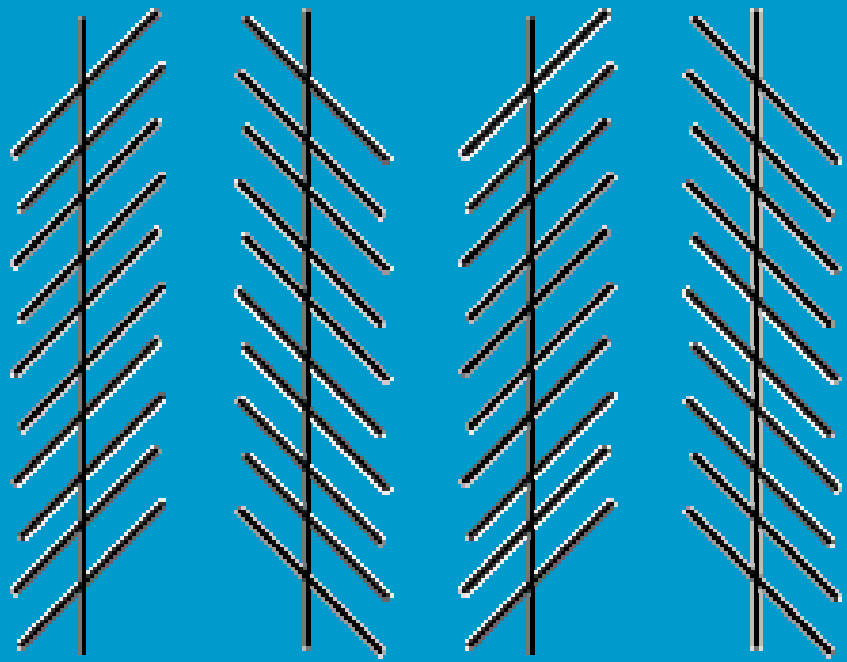
La spirale di Fraser

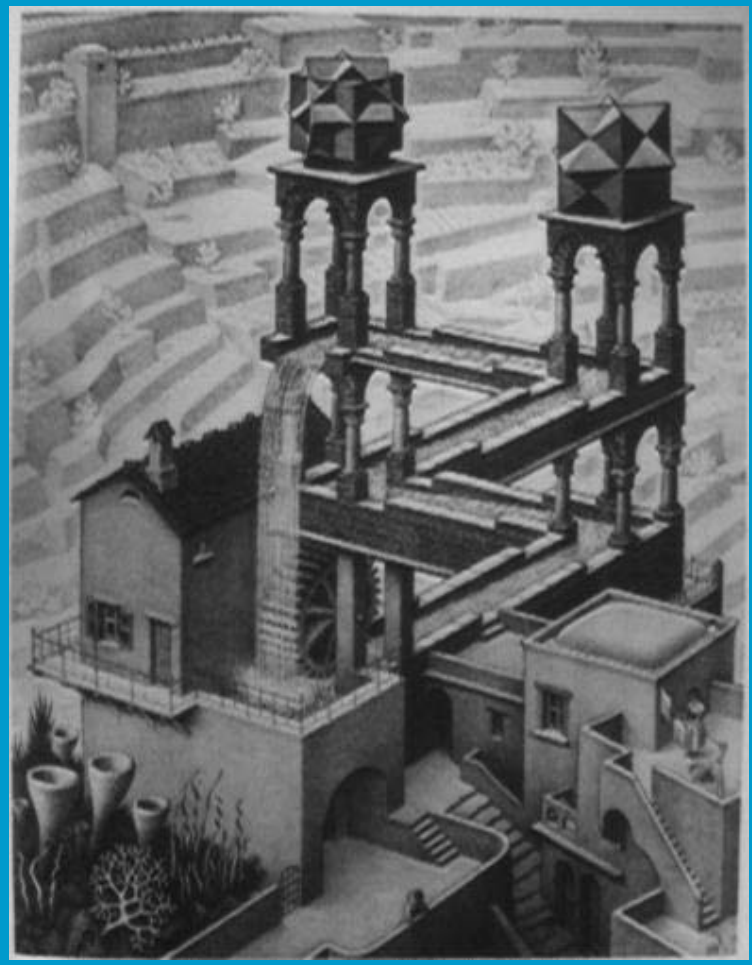


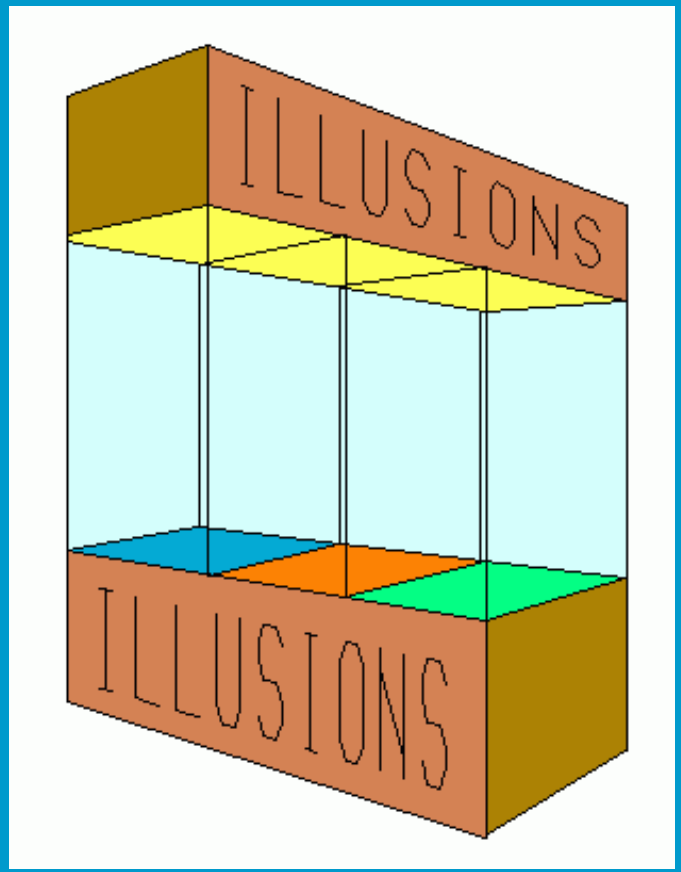
La griglia di Hermann-Hering



Illusione di Zoellner






ILLUSIONS
ILLUSIONS

Bottom-up e top-down

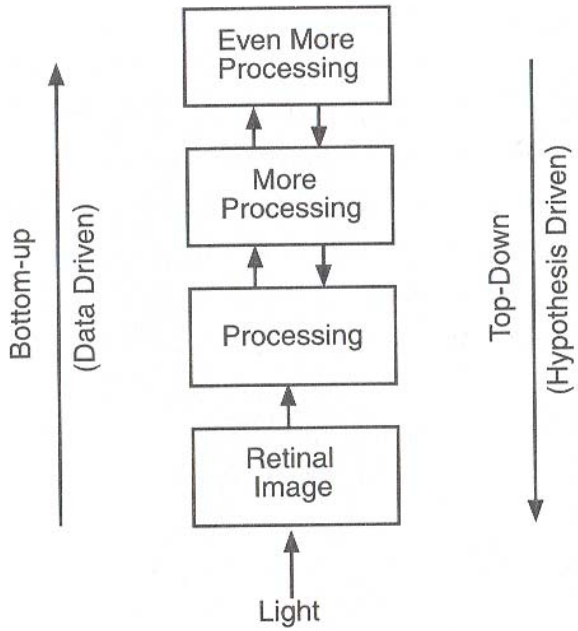
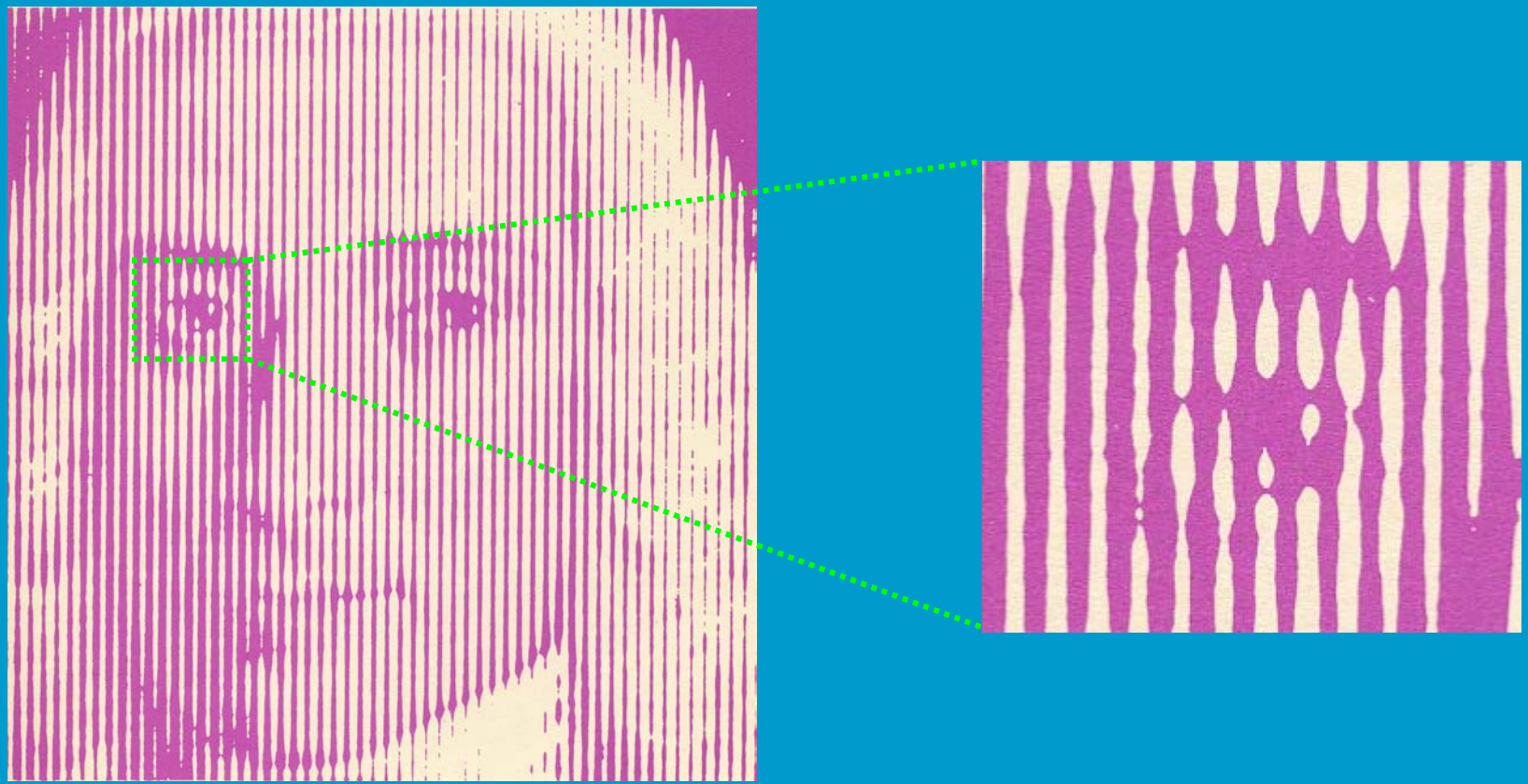


Figure 2.3.11 Bottom-up versus top-down processing. The two directions of processing are referred to as *bottom-up* (or *data driven*) from lower to higher levels of processing and *top-down* (or *hypothesis driven*) from higher to lower levels of processing.

Top-down o bottom-up?

Top-down o bottom-up?

Top-down o bottom-up?



La maschera di Chaplin

Figure 1. Photographs of a rotated hollow mask: (a) and (b) (black hat) show the front and side truly convex view; (d) (white hat) shows the inside of the mask; it appears convex although it is truly hollow; (c) is curiously confusing as part of the hollow inside is seen as convex, combined with the truly convex face. This is even more striking with the actual rotating mask. Viewing the hollow mask with both eyes it appeal's convex, until viewed from as close as a metre or so. Top-down knowledge of faces is pitted against bottom-up signalled information. The face reverses each time a critical viewing distance is passed, as 'downwards' knowledge or 'upwards' signals win. (This allows comparison of signals against knowledge by nulling.)

Contenuto del corso

- Rilevamento di feature (edge/corner detection)
- Segmentazione e raggruppamento percettivo
- Visione stereoscopica
- Shape-from-shading
- Riconoscimento di oggetti

Riferimenti bibliografici

- R. Jain, R. Kasturi, B. G. Schunck. *Machine Vision*. McGraw-Hill, 1995.
- E. Trucco, A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice-Hall, 1998.
- D. Forsyth, J. Ponce. *Computer Vision: A Modern Approach*. Prentice-Hall, 2003.
- Dispense e articoli vari…